MONTHLY WEATHER REVIEW

Editor, ALFRED J. HENRY

VOL. 57, No. 5
W. B. No. 985

MAY, 1929

CLOSED JULY 3, 1929
ISSUED AUGUST 30, 1929

CORRELATION BETWEEN RAINFALL AND RUN-OFF

551.578.1 : 551.482

By J. W. SHUMAN, Mem. Amer. Soc. C. E.

The problem of determining the relationship between rainfall and river run-off has engaged the attention of many investigators (1). Attempts have been made to derive equations that would enable one to compute the run-off from the amount of rainfall on the drainage basin. Our great need for some knowledge of this relationship is brought about due to the inadequate length of most river-gaging records. Rainfall data are generally available for a long term of years, whereas river gaging records usually cover only brief periods, 4 to 10 years. The uncertainties of a solution of this problem are acknowledged by every investigator and by every book on hydrology. The methods and equations that are given can only be used as a rough guide, and no attempt is made to get very far into the problem or to show the extent of the fitness of the proposed method.

In considering this subject, we are faced at the outset with raw data that are admittedly inaccurate. It is impossible to place enough rain-gages over a drainage area to guarantee an absolute measure of the amount of precipitation. Also, the best methods in vogue for determining the quantity of the "sec.-feet" run-off will not give an absolute measure. True, if the river be a mere stream and a good wier can be thrown across it, fairly accurate results can be secured; but, in general, reliance is placed in floats and velocity meters. Even with good meters, when the uncertainties of securing the true section of the stream, are taken into account, as well as the changing conditions of the river bed or the control, it is easy to understand why good engineers make some reductions in their figures, in order to be on the safe side. In other words, the accuracy of the data is recognized as uncertain.

Because no dependable solution of this problem has been put forth, hydrologists have been obliged to resort to the scheme of taking what river run-off records were available and making therefrom a duration curve, showing variation of discharge with percentage of time. They then might give some consideration to the amount of rain that fell during the period embraced by the duration curve data, as compared to the long-time mean, and this would be about all that they could do, unless some probable discharges were worked out on the probability theory.

The relationship between rainfall and run-off is evidently a very complicated one, and to date has defied the efforts of our ablest investigators to set down in mathematical terms just what it really is. It was noticed a long time ago that the calendar year run-off (January to January) did not always keep step with the current calendar year rainfall, and the "water year," October to October, was adopted as a better 12-months period to consider. But even with this change, not much was gained. It is very likely that we may never be able to derive an exact expression defining the relationship, but it is the hope of the writer in what follows to show some progress in this problem.

It is now two years since the writer first saw an article, by A. Streiff (2), that dealt with the various cycles in our weather elements. This was the starting impulse that led to the special studies in rainfall as outlined below. Instead of inspecting columns of rainfall data, typed on a sheet of paper, the data were studied by preparing a graph, plotting the values as ordinates against time as abscissa, and connecting the points by a smooth curve. Run-off was treated the same way, which led eventually to showing the probable fact run-off *is* proportional to rainfall but *lags* behind it in time.

Inasmuch as the ideas exploited herewith all hinge upon this lag, it is advisable to promptly define exactly what is meant. By *calendar-year* run-off, or rainfall, reference is made to the amount that discharged or fell for the period from January 1 to January 1. In plotting the magnitude of these values, no consideration is given as to what date in the year these plotted points are assigned. Considering either rainfall or run-off on the graph, the plotted point simply represents the magnitude of the value. Thus a smooth curve drawn through the various plotted points is considered to have no significance between two adjacent points. By the term *rainfall year*, reference is made to a 12-months period of rainfall that does not necessarily coincide with the calendar year period. As to the lag, I find that the calendar-year run-off is proportional, *in a higher degree*, to a 12-months rainfall period, if the latter be advance one or more months, than it is to the current calendar-year rainfall; viz, if the run-off lags four months behind the rainfall, it is meant that the calendar-year run-off is proportional to the rainfall from September 1, of the previous year, to September 1 of the current year.

Attempts to discover any workable relationship between the variables we are considering, using periods of time less than one year, have been fruitless, so that I have considered in the following studies only the annual values. The nature of this lag will be disclosed in a study of the Tennessee, Yadkin (N. C.), and Root (Minn.), rivers, and of the rainfall and run-off data of the Wagon Wheel Gap (Colo.) experiment. It is convenient to think of this run-off lag as analogous to electrical inductive effect. In a given circuit, through which an alternating current is being sent, by proper insertion in the circuit of inductance or capacity, the current may be made to either lag or lead, with respect to the impressed voltage, high inductance inducing a large lag. In investigating various sized streams, this lag seems to vary from 3 to 6 and even 9 months.

Acceptance of the fact of run-off lag requires thinking of the regimen of a stream in a modified way. Of course the rainfall that occurs after the rainfall year has ended, contributes to the calendar year run-off. It must be admitted that a lapse of time must take place before run-off can appear. We think of this time as relatively short, because streams usually respond at once to pronounced rainfall. However, this response is not the lag, as this momentary response has but little if any effect on the stream's mean annual flow. If a given section of the country has sufficient rainfall, and the slope of the land is proper, a river will eventually result, but this will not occur with a series of spasmodic rainstorms, even though some of them be very severe. In order that a river may be one in fact it requires the presence of water that has fallen some time back.

RAINFALL YEAR AND LOCATING THE RUN-OFF LAG

As stated before, the rainfall year is a 12-months period. For a 2-months lag of run-off, the corresponding rainfall year would be from November 1 of the previous run-off year to November 1, of the current year; for 6-months lag, it would be from July 1 to July 1, etc. Care should be taken to secure as many rainfall stations as possible over the drainage area of the river. If a large enough number of such stations is available, it will not be necessary to weight the records. For example, in the Tennessee River study below, the rainfall records were first weighted, and later used just as printed. There seemed to be very little discrepancy, the maximum being less than 2 per cent. Thus, by being able to use the records just as printed, a great deal of labor is saved.

In order to work most efficiently proper attention must be given to the scales employed in making the graphs. The writer plots the run-off values as ordinates one-half inch apart, and so chooses the vertical scale as to give the resulting smooth curve drawn through the plotted points considerable amplitude up and down. A graph is first made of the calendar year run-off. Next the rainfall records are prepared for use. Suppose one has 10 stations, and inspection shows that they are bona fide records (no estimates inserted, as is frequently the case). Suppose, further, that the run-off records cover a period only 6 years long, and ceased 6 years ago, and that we desire some index of what the run-off has been in the past 6 years. We prepare a blank page with the 10 stations listed at the left, one below the other, and opposite the monthly headings at the top of the page, we set down the monthly rainfalls for each of the 12 years. Each year will be made up of 10 stations, 12 months each. These monthly columns for each year are now averaged, and a fresh memo prepared, showing the mean rainfall for each month, for each year, set down in a vertical column.

The total calendar year rainfall is now secured for each year, and marked conveniently at the side of each year's column. Just for curiosity we might plot the calendar year totals of rainfall on a piece of thin paper, laid over the graph of the run-off, to see if the curves resemble each other. Very likely we will find here and there some resemblance, but it will most likely extend only over a year or two, and then the values will be opposite. We do not know what the lag is, but we start out with the assumption, say, that it is three months. A table is prepared of a 12-month rainfall period, running from October 1 to October 1, for each of the 12 years, and we try plotting again on a piece of thin paper laid over the run-off graph, setting down the first rainfall year value on the first run-off value ordinate. Various shifts of the rainfall year are thus made until we find the best synchronism of the curves. This sounds rather complicated, but after the work of preparing the rainfall data, as outlined above, is done, the method is quite simple and quick. If no synchronism can be detected, no matter what lag is given, the chances are that either an error has been made in the preparation of the rainfall data, or in securing the total or mean run-offs, or that the records themselves are inaccurate; for we must bear in mind the nature of the data with which we deal.

TENNESSEE RIVER

This river has been gaged since 1874, and the records are probably the most reliable of any stream in the United States. Footnotes to the data state that there was no artificial regulation of the flow prior to October 22, 1913; but that after that date the low flows might be affected by the Hales Bar lock and dam. This should be a good river with which to test this lag idea.

The discharge is taken as at Chattanooga, from Bulletin No. 34, Water Resources of Tennessee, issued by the State. The drainage area above Chattanooga is about 21,400 square miles. Eleven United States Weather Bureau station records were used—Chattanooga, Decatur, Charleston, Kingston, Clinton, Tazewell, Knoxville, Bryson City, Bluff City, Hot Springs, and Asheville. The mean calender-year run-off was computed from the records, and plotted in curve No. 1, Figure 1. The complete rainfall records for the above 11 stations are not available farther back than 1907. Table 1 gives the mean annual run-off and the 12 months rainfall over the basin for a 3½ months lag, as it was found that this lag brought about the highest state of agreement between the curves. These data are given in columns 2 and 3, from 1908 to 1926. The rainfall values in column 3 are plotted in curve No. 2, Figure 1. The similarity of the curves may be noted as almost perfect with the exception of the year 1915–16.

TABLE 1.—*Showing observed mean annual discharge of Tennessee River at Chattanooga, Tenn.; annual rainfall over drainage basin; mean annual computed discharge, from equation. Second-feet = 879.2 × rainfall − 5809; and accuracy of results*

Year (1)	Observed second-feet (2)	Rainfall [1] (3)	Computed second-feet (4)	Per cent column 4 is of column 2 (5)
1908	39, 366	47. 05	35, 557	90. 5
1909	47, 342	60. 31	47, 216	99. 8
1910	29, 185	43. 71	32, 621	111. 8
1911	34, 758	45. 66	34, 335	98. 8
1912	40, 175	*56. 73*	*44, 059*	*109. 7*
1913	34, 882	47. 19	35, 680	102. 2
1914	27, 512	40. 05	29, 403	106. 8
1915	37, 058	48. 89	37, 175	100. 2
1916	38, 691	54. 48	42, 090	108. 8
1917	43, 292	57. 06	44, 358	101. 7
1918	36, 233	42. 11	31, 214	86. 2
1919	37, 242	50. 49	38, 582	103. 5
1920	50, 183	58. 96	46, 029	91. 9
1921	35, 050	44. 89	33, 658	96. 0
1922	43, 003	50. 64	38, 714	90. 1
1923	41, 249	53. 87	41. 554	100. 6
1924	35, 583	48. 52	36, 850	103. 7
1925	23, 110	34. 72	24, 717	107. 0
1926	31, 294	47. 57	36, 014	115. 2

[1] 3½ months ahead of run-off.

In order to secure a measure of how close is the correlation of the curves thus depicted, we compute the correlation coefficient as given in Table 2. We have dealt here with data from 1908 to 1925, so as to compute the 1926 discharge.

A correlation coefficient of +0.9112 is found, indicating a high state of agreement. Assuming linear regression, and using mean values, we secure the regression equations for both second-feet and rainfall,

(I) Second-feet = 879.2 × rainfall − 5,809

(II) Rainfall = 0.000944 × second-feet + 13.83.

Using equation (I), we have computed the discharges for the period given in Table 1, and give same in column 4. The percentages of the computed to the observed values are given in column 5. The greatest deviation shown is 15.2 per cent, for 1926. Exclusive of this date, the computed values run remarkably close to the observed.

The trend of flow[1] in this well-gaged river is shown in curve A, Figure 1, from 1875 to 1926, and the Brückner cycle is plainly visible in the graph, as suggested by the dotted line. Streiff (**2**), (**3**), has shown how the run-off can be estimated in the future for a river of this class, with the available discharge records to work with.

[1] The trend of flow is found by drawing a median line through the graph; in other words, it is a form of smoothing.—ED.

TABLE 2.—*Correlation between rainfall and run-off, Tennessee River*

Year	Run-off—X	Rain-fall—Y	x	y	x^2	y^2	xy +	xy −
1908	39, 366	47. 05	+1, 927	−2. 14	3, 713, 300	4. 58	----------	4, 123. 8
1909	47, 342	60. 31	+9, 903	+11. 12	98, 070, 000	123. 65	110, 121. 4	----------
1910	29, 185	43. 71	−8, 254	−5. 48	68, 127, 000	30. 03	45, 231. 9	----------
1911	34, 758	45. 66	−2, 681	−3. 53	7, 188, 000	12. 46	9, 468. 9	----------
1912	40, 175	56. 73	+2, 736	+7. 54	7, 486, 000	56. 85	20, 629. 4	----------
1913	34, 882	47. 19	−2, 557	−2. 00	6, 538, 000	4. 00	5, 114. 0	----------
1914	27, 512	40. 05	−9, 927	−9. 14	98, 545, 000	83. 54	90, 732. 8	----------
1915	37, 058	48. 89	−381	−. 30	145, 161	. 09	114. 3	----------
1916	38, 691	54. 48	+1, 252	+5. 29	1, 567, 500	27. 98	6, 623. 1	----------
1917	43, 292	57. 06	+5, 853	+7. 87	34, 259, 000	61. 94	46, 063. 1	----------
1918	36, 233	42. 11	−1, 206	−7. 08	1, 454, 500	50. 13	8, 538. 5	----------
1919	37, 242	50. 49	−197	+1. 30	38, 809	1. 69	----------	256. 1
1920	50, 183	58. 96	+12, 744	+9. 77	162, 410, 000	95. 45	124, 508. 9	----------
1921	35, 050	44. 89	−2, 389	−4. 30	5, 707, 400	18. 49	10, 272. 7	----------
1922	43, 003	50. 64	+5, 564	+1. 45	30, 958, 000	2. 10	8, 067. 8	----------
1923	41, 249	53. 87	+3, 810	+4. 68	14, 185, 000	21. 90	17, 830. 8	----------
1924	35, 583	48. 52	−1, 856	−. 67	3, 444, 800	. 45	1, 243. 5	----------
1925	23, 110	34. 72	−14, 329	−14. 47	205, 330, 000	209. 38	207, 340. 6	----------
Means	37, 439	49. 19	----------	----------	41, 620, 415	44. 71	711, 896. 7 −4, 379. 9 +707, 516. 8	4, 879. 9 ----------

$$(41{,}620{,}415)^{\frac{1}{2}} = 6{,}451.3 = \sigma_x \qquad r_{xy} = \frac{+707{,}516.8}{18 \times 6{,}451.3 \times 6.686} = +.9112$$

$$(44.71)^{\frac{1}{2}} = 6.686 = \sigma_y \qquad b_{xy} = \frac{0.9112 \times 6{,}451.3}{6.686} = 879.2$$

$$b_{yx} = \frac{0.9112 \times 6.686}{6{,}451.3} = 0.000.944$$

Regression equations:
(I) Second-feet = 879.2 × rainfall − 5,809
(II) Rainfall = 0.0009444 × second-feet + 13.83

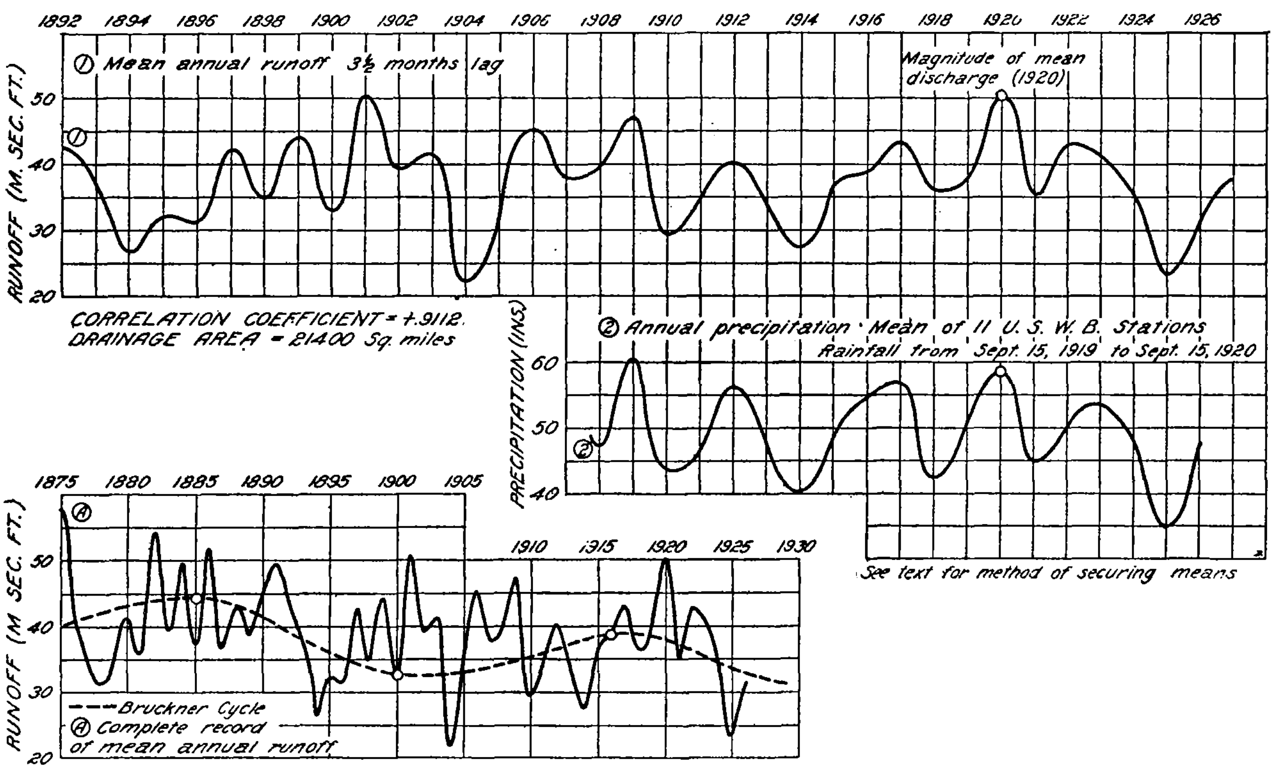

FIGURE 1.—Tennessee River, precipitation and run-off graph

YADKIN (N. C.) RIVER

This river has a drainage area above Salisbury, N. C., of about 3,400 square miles, and the gaging records at Salisbury are supposed to be of fair accuracy, with no ice troubles. Records are available in the United States Geological Survey's Water Supply Bulletins, from 1903 to 1909, inclusive, for the year 1912, and from 1914 to 1919, inclusive. There are gaps in the record for 1910 and 1911 and for 1913. What will this method do for us here? The records are of scarcely continuous enough length to utilize Streiff's method. Rainfall records are used from Salisbury, Settle, Statesville, Winston-Salem, Brewers, Lenoir, and Mount Airy—seven in all. By the process outlined, the lag of run-off behind rainfall in this case is found to be about four months. Table 3 lists the data.

TABLE 3.—*Showing observed mean annual discharge of Yadkin River, near Salisbury, N. C.; annual rainfall; computed mean annual discharge from regression equation, second-feet = 161.4 × rainfall − 2964.; and accuracy of results*

Year	Observed second-feet	Rainfall [1]	Computed second-feet	Per cent column 4 is of column 2
(1)	(2)	(3)	(4)	(5)
1903	6,850	58.77	6,521	95.2
1904	2,873	39.69	3,442	119.7
1905	4,539	48.90	4,928	108.6
1906	6,620	56.73	6,192	93.6
1907	4,480	40.34	3,547	79.2
1908	6,480	57.61	6,334	96.1
1909	5,710	53.53	5,676	94.4
1910		43.38	4,038	
1911		37.30	3,056	
1912	4,950	54.69	5,863	118.5
1913		52.34	5,484	

[1] 4 months ahead of run-off.

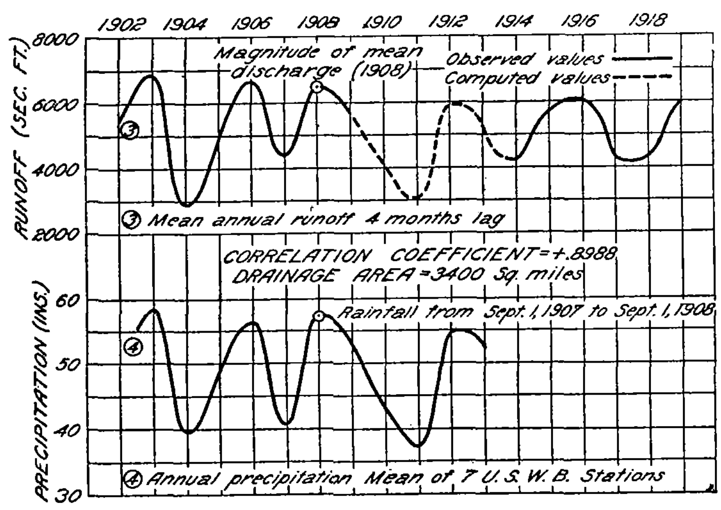

FIGURE 2.—Yadkin River, precipitation and run-off graph

Table 4 gives the computations for the correlation coefficient between the rainfall and run-off data, which has been plotted in curves No. 3 and 4 in Figure 2. We have taken the data only up to and inclusive of the year 1913, although doubtlessly the correlation would have been closer had we included the balance of the record. However, we wish merely to show the method, and as it is, we secure a correlation coefficient of +0.8988. The regression equations are:

(I) Second-feet = 161.4 × rainfall − 2964.
(II) Rainfall = 0.005 × second-feet + 25.71.

Using equation (I) we compute the discharge in second-feet, and thus fill in the gaps. Table 3 shows the computed values in column 4, and their accuracy with respect to the observed values in column 5. Here we note a maximum discrepancy of 19.7 per cent. This may apply to the gap computed values, too; but nevertheless, our computed values are better than wild guesses. Had we taken the complete record up to 1919, we probably would have secured a better fit.

TABLE 4.—*Correlation between rainfall and run-off, Yadkin River*

Year	Run-off—X	Rain-fall—Y	x	y	x^2	y^2	xy +	xy −
1903	6,850	58.77	+1,537	+7.49	2,362,369	56.10	11,512.1	
1904	2,873	39.69	−2,440	−11.59	5,953,600	134.33	28,279.6	
1905	4,539	48.90	−774	−2.38	599,076	5.66	1,842.1	
1906	6,620	56.73	+1,307	+5.45	1,708,249	29.70	7,123.2	
1907	4,480	40.34	−833	−10.94	693,889	119.68	9,113.0	
1908	6,480	57.61	+1,167	+6.33	1,361,889	40.07	7,387.1	
1909	5,710	53.53	+397	+2.25	157,609	5.06	893.3	
1912	4,950	54.69	−363	+3.41	131,769	11.63		1,237.8
Mean	5,313	51.28			1,621,056	50.28	66,150.4 −1,237.8 +64,912.6	1,237.8

$$(1{,}621{,}056)^{\frac{1}{2}} = 1{,}273.2 = \sigma_x \qquad r_{xy} = \frac{64{,}912.6}{8 \times 1{,}273.2 \times 7.091} = +0.8988$$

$$(50.28)^{\frac{1}{2}} = 7.091 = \sigma_y \qquad b_{xy} = \frac{0.8988 \times 1{,}273.2}{7.091} = 161.4$$

$$\frac{b_{yx} = 0.8{,}988 \times 7.091}{1{,}273.2} = 0.005005$$

Regression equations:
(I) Second-feet = 161.4 × rainfall − 2964.
(II) Rainfall = 0.005 × second-feet + 25.71.

ROOT RIVER

This river has a drainage area, near Houston, Minn., of about 1,560 square miles. It is a small stream, and yet has important water power along its course. Unfortunately there are available only three years of complete monthly mean flows. These even are in doubt, due to hard freeze-ups in winter. We have taken the

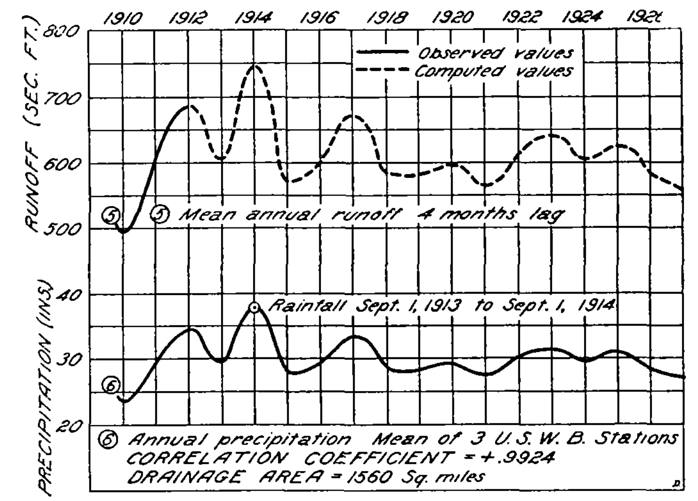

FIGURE 3.—Root River, precipitation and run-off graph

mean annual discharge for the years 1910, 1911, and 1912 as a mean of the United States Geological Survey records and of Adolph F. Meyer's computed results (from his book on hydrology). This is all we have to work with, if we would like some idea of the discharge variation in the following years.

Rainfall data were averaged from Caledonia, Grand Meadow, Rochester, and Chatfield, Minn. The run-off lag here was again found to be about four months. In Figure 3, curve No. 5 shows the run-off and curve No. 6 the rainfall graphs. Of course with but three records to consider, we naturally expect to find a high correlation, and are not surprised when the computations in Table 5 disclose a figure of +0.9924. The computed values are shown in Table 6, and naturally for the three known years, the values are very close to observed. As for the balance of the time, we can only assume an accuracy of say 10 to 15 per cent for the computed values from 1913 to 1927. The data on which we have based our relations is meager, but if we reflect upon the results obtained with the Tennesesee River, we can have greater confidence in the results we have secured. Again, it is a great deal closer to the actual discharge than any guess.

TABLE 5.—*Correlation between rainfall and run-off, Root River*

Year	Run-off—X	Rain-fall—Y	x	y	x^2	y^2	xy +	xy −
1910	493	23.5	−104	−5.9	10,816	34.81	613.6	
1911	610	29.9	+13	+.5	169	.25	6.5	
1912	689	34.8	+92	+5.4	8,464	29.16	496.8	
Means	597	29.4			6,483	21.41	1,116.9	0.00

$$(6{,}483)^{\frac{1}{2}}=80.52=\sigma_x \qquad r_{xy}=\frac{1116.9}{3\times 80.52\times 4.627}=+0.9924$$

$$(21.41)^{\frac{1}{2}}=4.627=\sigma_y \qquad b_{xy}=\frac{.9924\times 80.52}{4.627}=17.39$$

Regression equation for discharge:
(I) Second-feet = 17.39 × rainfall + 86.

TABLE 6.—*Showing observed mean annual discharge of Root River near Houston, Minn.; rainfall, and computed values from equation: second-feet = 17.39 × rainfall + 86*

(1) Year	(2) Observed second-feet	(3) Rainfall [1]	(4) Computed second-feet	(1) Year	(2) Observed second-feet	(3) Rainfall [1]	(4) Computed second-feet
1910	493	23.5	495	1919		28.5	582
1911	610	29.9	606	1920		29.5	599
1912	689	34.8	691	1921		27.8	569
1913		29.6	605	1922		30.3	613
1914		38.1	749	1923		31.9	641
1915		28.0	573	1924		29.9	606
1916		29.7	602	1925		31.2	629
1917		34.0	677	1926		28.8	584
1918		28.5	582	1927		27.0	556

[1] 4 months ahead of run-off.

WAGON WHEEL GAP (COLO.) EXPERIMENT RAINFALL AND RUN-OFF DATA

The data used in discussing the previous examples might be open to question as to accuracy—in fact it is ordinary rainfall and run-off data as collected and published. Using such data and obtaining such relatively high correlation coefficients might be challenged as to results on the grounds of pure accident. The Wagon Wheel Gap data, however, are beyond dispute very accurate. For the conditions obtaining in the securing of this data, reference should be made to Supplement No. 30 of the MONTHLY WEATHER REVIEW.

With these precise data of rainfall and resulting run-off, if we can secure a high correlation coefficient, it should lend reality to the claims here set forth. We will consider watershed A, as this was left untouched during the period of the experiment, whereas the B watershed was denuded after half of the time had expired. Watershed A has an area of about 222 acres, the run-off consisting of a streamlet fed by several springs, and measured by means of a V-notch wier, carefully calibrated. Table 7 gives, for the period 1912 to 1925, inclusive, the total run-off in inches over the area; the calendar year and the July 1 to July 1 year rainfall; the computed run-off from the regression equation derived, and the accuracy of the results.

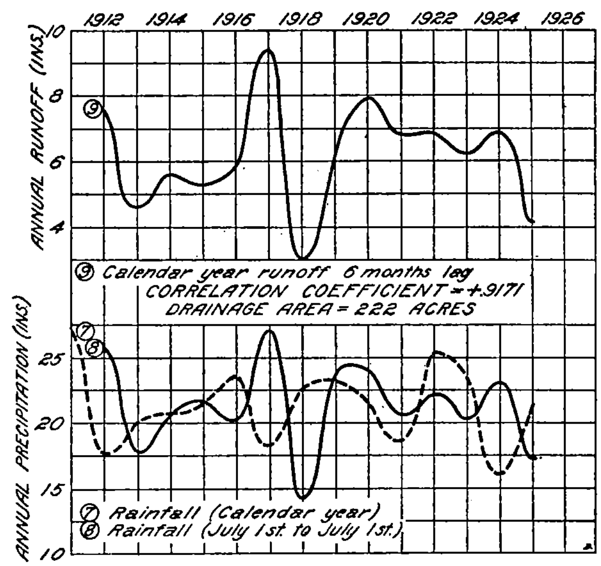

FIGURE 4.—Wagon Wheel Gap, Colo., graph of rainfall and run-off

Figure 4 shows in curve No. 9 the run-off graph. The dotted curve, No. 7, is the calendar-year rainfall graph, and it is seen that there is but little resemblance between it and the run-off. Curve No. 8 is the July 1 to July 1 rainfall year graph, this particular period being found by repeated trial (the same as with the other examples quoted). There seems to be a remarkably close relationship between curves 8 and 9, and the correlation coefficient, computed in Table 8 discloses this to be +0.9171. The measuring of the rainfall and run-off was precisely done, but the area involved was very small.—From Supplement No. 30, MONTHLY WEATHER REVIEW.

TABLE 7.—*Showing run-off in total inches over Watershed A, Wagon Wheel Gap, Colo.; rainfall for both calendar year and July 1 to July 1 year; computed results and accuracy*

Year	Observed run-off	Rainfall		Computed run-off	Per cent column 5 is of column 2
		July 1 to July 1	Calendar year		
(1)	(2)	(3)	(4)	(5)	(6)
1912	7.57	25.67	17.94	8.02	106.3
1913	4.64	17.86	20.22	4.67	100.6
1914	5.61	20.84	20.88	5.95	106.0
1915	5.33	21.94	21.56	6.42	120.0
1916	5.87	20.04	23.62	5.61	95.6
1917	9.47	27.25	18.44	8.71	92.0
1918	3.04	14.19	22.87	3.09	101.6
1919	6.18	24.29	23.31	7.43	120.3
1920	7.97	24.20	21.64	7.40	92.8
1921	6.91	20.79	18.52	5.93	85.8
1922	6.84	22.23	25.38	6.55	95.7
1923	6.23	20.06	23.42	5.62	90.3
1924	6.95	23.37	16.00	7.04	101.3
1925	4.15	17.32	21.34	4.44	106.9

TABLE 8.—*Correlation coefficient between rainfall and run-off, Wagon Wheel Gap experiment, Watershed A*

Year	Run-off— X	Rain-fall— Y	x	y	x^2	y^2	xy +	xy −
1912	7.57	25.67	+1.37	+4.24	1.88	17.98	5.81	
1913	4.64	17.86	−1.56	−3.57	2.43	12.74	5.57	
1914	5.61	20.84	−.59	−.59	.35	.35	.35	
1915	5.33	21.94	−.87	+.51	.76	.26		0.44
1916	5.87	20.04	−.33	−1.39	.11	1.93	.46	
1917	9.47	27.25	+3.27	+5.82	10.69	33.87	19.03	
1918	3.04	14.19	−3.16	−7.24	9.99	52.42	22.88	
1919	6.18	24.29	−.02	+2.86	0	8.17		.06
1920	7.97	24.20	+1.77	+2.77	3.13	7.67	4.90	
1921	6.91	20.79	+.71	−.64	.50	.41		.45
1922	6.84	22.23	+.64	+.80	.41	.64	.51	
1923	6.23	20.06	+.03	−1.37	0	1.88		.04
1924	6.95	23.37	+.75	+1.94	.56	3.76	1.46	
1925	4.15	17.32	−2.05	−4.11	4.20	16.89	8.43	
Mean	6.20	21.43			2.50	11.355	69.40	.99
							−.99	
							+68.41	

$$(2.50)^{\frac{1}{2}}=1.58=\sigma_x \qquad r_{xy}=\frac{+68.41}{14\times1.58\times3.37}=+0.9174$$

$$(11.355)^{\frac{1}{2}}=3.370=\sigma_y \qquad b_{xy}=\frac{0.9174\times1.58}{3.37}=0.430$$

$$b_{yx}=\frac{0.9174\times3.37}{1.58}=1.955$$

Regression equations:
(I) Run-off=0.43×rainfall−3.01.
(II) Rainfall=1.955×run-off+9.31.

CONCLUSIONS

The subject matter discussed is the result of special study, in the search of a better understanding of the problem. Having constant use for rainfall and run-off data, the writer has felt that the work of previous investigators has carried a solution only part way. It is contended that their difficulty has lain partly in the fact that the nature of the lag of run-off behind rainfall has not been sufficiently considered. A higher state of agreement seems to result, in every case that has been tried by the writer, *when* some lag is given to the run-off. The amount of this lag, as given in the examples cited, may be open to greater refinement and may be derived by more elegant methods. It may be given different values and still not vary the average results of computed values very much. For example, in the case of the Yadkin River, the lag of four months was adopted and the correlation coefficient found, considering the years 1903 to 1909, inclusive, and the year 1912. The coefficient was +0.8988. If now, in addition to the years above, we take the years 1914 to 1919, inclusive, making 14 years in all, and take a lag of 2½ months only, the correlation coefficient comes out at +0.9061 (about the same figure, although slightly higher); and the computed results differ very little from those in Table 3 (the average). However, if *no* lag is given the run-off, the correlation is very small.

Variation in temperature and wind movement must have considerable effect upon the amount of rainfall that is responsible for the mainstay of the stream's flow. Attempts to include consideration of these factors in the Wagon Wheel Gap data have led nowhere up to the present. Run-off seems to be roughly inversely proportional to wind movement and temperature, and these two latter variables display considerable change in value in the Wagon Wheel Gap data. Their influence may be the missing link of some 11 to 12 per cent that we seem to lack in the correlation coefficient. It is the hope of the writer that this article may provoke further study leading to a complete and elegant solution.

LITERATURE CITED

(1) Elements of Hydrology, by A. F. Meyer. John Wiley & Sons, 1917.
Hydrology, by D. W. Mead. McGraw-Hill Book Co., 1919.
Failure of Hydraulic Project from Lack of Water Prevented by Better Hydrology, by R. E. Horton. Eng. News-Record, vol. 78, p. 490, 1917.
Comparison Between Rainfall and Run-off in Northern United States, by J. C. Hoyt. Trans. Am. Soc. C. E., vol. LIX, 1907.
Water Supply of New Jersey, by C. C. Vermeule, 1894.
Relation of Rainfall to Run-off, by G. W. Rafter. U. S. Water Supply Paper No. 80. 1903.
Derivation of Run-off from Rainfall Data, by Joel D. Justin, Trans. Amer. Soc. C. E., vol. LXXVII, p. 346, 1914.
(2) On the Investigation of Cycles and the Relation of the Bruckner and Solar Cycles, by A. Streiff. MONTHLY WEATHER REVIEW, July, 1926. Washington.
(3) Notes on Estimating Run-off, by A. Streiff. MONTHLY WEATHER REVIEW, March, 1928. Washington.

RAINFALL PERSISTENCY AT SAN JUAN, P. R.

551.578.1 (729.5)

By C. L. RAY

[Weather Bureau, San Juan, P. R., April 22, 1929]

The following notes relating to rainfall persistency at San Juan, P. R. were suggested by an original study of Besson, an abstract of which appeared in the MONTHLY WEATHER REVIEW, June, 1924, page 308,[1] "On the Probability of Rain (at Montsouris near Paris)." The factor of persistency or the tendency of rainfall to repeat itself for 1, 2, 3, or a greater number of consecutive days, was found to be markedly higher than the general probability, or in other words, the probability independent of what took place the day before. For example, at Monsouris the general probability for rain is 0.525. In a 50-year record there were 9,580 rainy days out of a possible 18,261, while there is an increase of 18 per cent or to 0.704 in the probability of a continuation of rain on a second day. There is also a gradual increase in the probability up to and including 15 days of rain and except for a short lapse a further increase for the higher groups. A similar study by Blair[1] based upon 30 years' records at Lincoln, Nebr., shows much the same effect of the persistency factor. At this station a general probability of 0.394—4,312 days of rain out of 10,956 possible is followed by an increase to 0.540 for a second day, or 14 per cent, and a gradual rise in values in consonance with the increase of consecutive days of rain.

In the present paper we have taken as a basis the 29-year-record at San Juan, P. R., comprising 6,205 rainy days out of 10,585 possible, for a general probability of 0.586. In Table 1 are given the total of single or separated days of rain, and the number of groups of 2, 3, 4, 5, etc., consecutive days upon which precipitation occurred. In the same table is shown the calculated days of rain (the number that would be expected) derived from the equation of general probability. The actual cases with rain exceed the calculated where more than 5 days are considered but are less where 1, 2, 3, 4, or 5 days is the factor. Thus we have a first indication of the part played by a markedly long series of daily showers at San Juan, which as

[1] Besson: Vol. 52, Mo. Wea. Rev., p. 309. Blair: Vol. 52, Mo. Wea. Rev., p. 350.